Asymptotic extremal distribution for non-stationary, sontrgly-dependent data

Gonzalo Perera [1] and Carolina Crisci [1]

[1] Polo de Desarrollo Universitario Modelizacion y Analisis de Recursos Naturales, Centro Universitario Regional del Este, Universidad de la Republica. Interseccion de Ruta 9 con Ruta 15, Rocha, Uruguay.

1. Abstract

2. Introduction

The statistical analysis of extreme values has a wide and vast domain of applications on many disciplines. Extreme wind speeds are a key input for design in Structural Engineering, maximun levels of traffic are crucial for design and operation in Telecommunications Networks, maximun tides are essential for any policy concerning coast resources mangement, extreme events on chemical, physical or biological conditions may affect dramatically very sensitive ecosystems, etc.([2], [4]. [5], [8],[9],[10], [17]).

The classical Fisher-Tippet-Gnedenko theory determines the asymptotic behavior of the maximum of an iid sequence of random variables, with no other possible non-degenerated limit that the class of extremal distributions ([6],[7]).

Since three decades ago we know that classical theory applies to a stationary and weakly dependent secquence of random variables, thanks to the work of Leadbetter, Lindgreen and Rootzen ([12],[13], [14], [18], [19], [20], [21], [22]).

However in many real examples, a very big collection of maximal registers of a large series of measurements, does not fit to any extremal distribution. This often is associated to phenomena where the observed system may assume different states that produces drastical changes and that even may introduce long range dependence on data. Think, for instance, of the classical series of data of the Nile River with its very ancient regimen of anual floods, and other large series of hydrological data, in particular those related to the impact of climate change ([11], [15], [23]).

At a theoretical level, we will apply in the extremal context a method that has been used for different purposes in presence of non-stationary and possibly strong dependent data, where a random covariable indicates the state of the system, and the global behavior may be represented by a mixture of models. For instance, that is the case of Compound-Poisson approximation of High-Level Exceedances of time series ([3]), asymptotic behaviour of averages ([16]), and Nadaraya-Watson regression for functional data ([1]).

In this paper we will consider data that depend of two independent components: on hone hand, a categorical covariable process that describe the state of the system, that may be neither stationary nor weakly dependent, and that only satisfies that the mean frequency of each state has a (possibly random) limit, and, on the other hand, an iid noise. We will assume that for a given state j on the covariable, the maximal asymptotic distribution is non-degenerated and depends on j. In the main result of the paper we will show that if the covariable process satisfies a weak-dependence assumption, the maximal asymptotic distribution of our data is still an extremal distribution, consistently with Lindgren, Leadbetter and Rootzen results. But we will also show that if this covariable process has a strong dependence structure, then this maximal assymototic distribution is no longer an extremal one, but a mixture of extremal distributions.

The paper is organized as follows:

We begin in Section 2 by a very brief summary of classical Fisher-Tippet-Gnedenko theory, in particular the characterization of the Maximal Domains of Attraction of extremal distributions and an elementary Lemma that we will use later.

We then present mi Section 3 our model, its hypotheses, some examples and the main results with some corollaries.

In section 4 we will fit some simulated strong dependent data to a mixture of extremal distributions, showing that they do not fit extremal distributions. In particular, we will show the impact on return times and return values of missfitting data to a single extremal distribution.

Finally in section 5, we present the conclussions and some further work in progress.

3. Brief review of classical extreme values theory

The classical Fisher-Tippet-Gnedenko theory states that a distribution F belongs to the Maximal Domain of Attraction of an extremal distribution (Weibull, Gumbel, Frechet) H if there exists an iid sequence $X_1, \ldots, X_n, \ldots$ of random variables with distribution F such that for some deterministic sequences d_n and $c_n > 0$, we have that

$$\frac{max(X_1,\ldots,X_n)-d_n}{c_n}\underset{\mathrm{n}}{\overset{\mathrm{w}}{\longrightarrow}} H(\#eq:1)$$

Recall that $M_F = sup\{t\in\mathbb{R} : F(t)<1\}, F^{-1}(p) = inf\{x\in\mathbb{R} : F(x)\geq p\}\ \forall p\in(0,1)$.

Then the three Maximal Domains of Attraction (MAD) are fully described.

1. Frechet of order α: F belongs to the $MAD(\Phi_\alpha)$, where $\Phi_\alpha(x) = exp\{-x^{-\alpha}\}\ \forall x>0$ and $\alpha>0$ is called the order parameter, if and only if, $M_F=+\infty$ and for x tending to infinity, $1-F(x)$ is $\frac{L(x)}{x^{-\alpha}}$ for some slowly varying function L. In that case, in #eq:1, the deterministic sequences are $d_n=0, c_n=n^{1/\alpha}$.
2. Weibull of order α: F belongs to the $MAD(\Psi_\alpha)$, where $\Psi_\alpha(x)=exp\{-(-x)^\alpha\}\ \forall x<0$, if and only if $M_F<\infty$ and for x tending to M_F^-, $1-F(x)=(M_F-x)^\alpha L(\frac{1}{M_F-x})$ with L a slowly varying function. In this case $\alpha_n=M_F$ and $c_n=n^{\frac{-1}{\alpha}}$.
3. Gumbel: F belongs to the $MAD(\Lambda)$, where $\Lambda(x)=exp\{-e^{-x}\}\ \forall x\in\mathbb{R}\ iid$ and only if there exist an $a<M_F$ (which may be finite or infinite), some $c>o$ and a positive function h, with density h', such that $\lim_{x\to M_F^-} h'(x)=0$ and, for x tending to M_F^-, $1-F(x)$ is equivalent to $c\ exp\{-\int_a^x \frac{1}{h(t)}dt\}$, and $d_n=F^{-1}(1-1/n)$, $c_n=h(d_n)$.

$\underline{Lemma 1}$:

If we denote by $c_n(\Phi_\alpha)$, $c_n(\Psi_\alpha)$, $d_n(\Psi_\alpha)$,$c_n(\Lambda)$, $d_n(\Lambda)$, the deterministic sequences corresponding to each MDA, we then have

i. If $\alpha_1<\alpha_2$ $\frac{c_n(\Phi_{\alpha_2})}{c_n(\Phi_{\alpha_1})}\ \vec{n}\ 0$, $\frac{c_n(\Psi_\alpha)}{c_n(\Phi_{\alpha_1})}\ \vec{n}\ 0\ \forall\alpha>0$, $\frac{d_n(\Psi_\alpha)/c_n(\Psi_\alpha)}{c_n(\Phi_{\alpha_1})}\ \vec{n}\ 0\ \forall\alpha>0$, $\frac{c_n(\Lambda)}{c_n(\Phi_{\alpha_1})}\ \vec{n}\ 0$, $\frac{d_n(\Lambda)/c_n(\Lambda)}{c_n(\Phi_{\alpha_1})}\ \vec{n}\ 0$

ii. $\frac{c_n(\Psi_\alpha)}{c_n(\Lambda)}\ \vec{n}\ 0$, $\frac{d_n(\Psi_\alpha)/c_n(\Psi_\alpha)}{c_n(\Lambda)}\ \vec{n}\ 0$

iii. If $\alpha_1<\alpha_2$, $\frac{c_n(\Psi_{\alpha_1})}{c_n(\Psi_{\alpha_2})}\ \vec{n}\ 0$, $\frac{d_n(\Psi_{\alpha_1})/c_n(\Psi_{\alpha_1})}{c_n(\Psi_{\alpha_2})}\ \vec{n}\ 0$.

4. Main Results

We will assume that the process Y satisfies (H1). For any state $j=1,\ldots,k$, there exists a (possibly random) $b_j>0$ such that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}1\{Y_{i=j}\}=b_j\ a.s.$$

If $I(t)=\sigma\{Y_i : i\geq t\}$, and $I(\infty)$ = $\bigcap_{t>1}^{\infty} I(t)$, then, since $\forall j\ b_j$ is $I(\infty)$ -measurable,are deterministic, but, if $I(\infty)$ is not trivial, b_j may be non-deterministic, corresponding to strong dependence on the process Y. Let us show this in a very simple case.

$\underline{Example 1}$

Let U be a random variable such that $P(U=1)=p, P(U=2)=1-p$. Let $\sigma_1,\ldots,\sigma_n$,...an iid sequence of random variables on $\{1,2\}$ independent of U such that $P(\sigma_i(1)=1)=\delta$,$P(\sigma_i(1)=2)=1-\delta$,$P(\sigma_i(2)=1)=\eta$, $P(\sigma_i(2)=2)=1-\eta$, with $\sigma_i(1)$,$\sigma_i(2)$ independent, $0<\delta<1$,$0<\eta<1$.

Thus, $\frac{1}{n}\sum_{i=1}^n 1\{Y_i=1\}/U=1$ has the same distribution than $\frac{1}{n}\sum_{i=1}^n 1\{\sigma_i(1)=1\}\underset{\mathrm{n}}{\overset{\mathrm{a.s.}}{\longrightarrow}} P(\sigma_i(1)=1)$ by the Strong Law of Large Numbers $=\delta$

On the other hand $\frac{1}{n}\sum_{i=1}^n 1\{Y_i(1)=1\}/U=2$ has the same distribution than $\frac{1}{n}\sum_{i=1}^n 1\{\sigma_i(2)=1\}\underset{\mathrm{n}}{\overset{\mathrm{a.s.}}{\longrightarrow}}$ $P(\sigma_i(2)=1)=\eta$

Therefore if we assume that $\delta\neq\eta$, we have that

$$b_1=\begin{cases}\delta \text{ if } U=1\\ \eta \text{ if } U=2\end{cases}$$

b_1 is not-deterministic and $I(\infty)$ is not trivial. Similar treatment applies to b_2. Indeed, for the sake of simplicity we will assume (H2) $\forall b_j$ may only assume a finite number of values.

Our data will be of the form:

$$X_i=f(\varepsilon_i,Y_i)(\#eq:2)$$

where f is unknown we will assume (H3)

i. $\varepsilon_1,\ldots,\varepsilon_n,\ldots iid$

ii. $Y_1, \ldots, Y_n, \ldots$ satisfies (H1)

iii. The processes $\varepsilon_1, \ldots, \varepsilon_n, \ldots$ and $Y_1, \ldots, Y_n, \ldots$ are independent among them.

Finally we will assume that the state of the system is observable (speach aca o antes al introducir Y) and that it affects the extremal behavior of our data.

More precisely, we shall assume:

H4) There exists an integer $f, 1 < f < k$ and an integer $g, 1 < g, f + g < k$ such that

a. For $j = 1, \ldots, f$ the iid process $f(\varepsilon_1, j), \ldots, f(\varepsilon_n, j), \ldots$ belongs to $MDA(\Phi_{\alpha_j})$, where $0 < \alpha_1 < \ldots < \alpha_f$

b. For $j = f + 1, \ldots, f + g$ the iid process $f(\varepsilon_1, j), \ldots, f(\varepsilon_n, j), \ldots$ belongs to $MDA(\Lambda)$

c. For $j = f + g + 1, \ldots, k$, the iid process $f(\varepsilon_1, j), \ldots, f(\varepsilon_n, j), \ldots$ belongs to $MDA(\Psi_{\alpha j})$, with $0 < \alpha_{f_{f+g+1}} < \ldots < \alpha_k$

Then we have the following result

$\underline{Theorem 1}$:

Under (H1), (H2), (H3) there exists a random variable Z / $\frac{max(X_1,\ldots,X_n)}{n^{1/\alpha_1}} \underset{\text{n}}{\overset{\text{w}}{\longrightarrow}} Z$

In addition

a. If $I(\infty)$ is trivial, then, the distribution of Z is $F_z(x)$=$\sum_{i=1}^{k} p_i = 1\Phi_{\alpha_1} \frac{x}{b_1^{1/\alpha_1}}$

b. If $I(\infty)$ is not trivial and b_1 assumes the distribution of Z is $F_z(x) = \sum_{i=1}^{k} p_i \Phi_{\alpha_n}(\frac{x}{v_i} 1/\alpha_1)$

$\underline{Proof}$:

$$P(\frac{max(X_1, \ldots, X_n)}{n^{1/\alpha_1}} \leq x) = \int_{1,\ldots,k^\infty} P(\frac{max(X_1, \ldots, X_n)}{n^{1/\alpha_1}} \leq x)/Y_1 = j_1, \ldots, Y_n = j_n) dPY(j_1, \ldots, j_n, \ldots)(\#eq:3)$$

We will re-order the maximum according to the state j taken by each Y_i

Then #eq:3 equals to

$$\int_{\{1,\ldots,k\}^\infty} P(max\{(\frac{max\{X_{i_1}^1, \ldots, X_{i_n}^1\}}{n^{1/\alpha_1}} 1\{Y_{i_1}^1 = \ldots = Y_{i_{n_1}}^1 = 1\}), \ldots, (\frac{max\{X_{i_1}^k, \ldots, X_{i_{n_k}}^k\}}{n^{1/\alpha_1}} 1\{Y_{i_1}k \ldots = Y_{i_{n_k}}k = k\})\} \leq l/Y_n = j_1, \ldots, Y_n = j_n) dP^Y(j_1, \ldots, j_n, \ldots)(\#eq:4)$$

Among the blocks of maximum values taken with Y fixed on a given state, the MDA is different, and taking into account *Lemma1*, and (H4), all the blocks after the first, tends in probability to zero, and therefore, the limit as n tends to infinity of #eq:4 is the same than

$$\lim_n \int_{\{1,\ldots,k\}^\infty} P(\frac{max\{f(\varepsilon_{i_1}^1, 1), \ldots, f(\varepsilon_{n_1}^1, 1)\}}{n^{\alpha_1}} \leq x/Y_1 = j_1, \ldots, Y_n = j_n) dP^Y(j_1, \ldots, j_n, \ldots)(\#eq:5)$$

Using that ε_i are iid, the previous limit is the same than

$$\lim_n \int_{\{1,\ldots,k\}^\infty} P(\frac{max\{f(\varepsilon_1, 1), \ldots, f(\varepsilon_1, 1)\}}{n^{1/\alpha_1}} / \sum_{i=1}^{n} 1\{Y_i = 1\} = n_1) dP^Y(j_1, \ldots, j_n)$$

$$= \lim_n \int_{\{1,\ldots,k\}^\infty} \frac{P(max\{f(\varepsilon_1, 1), \ldots, f(\varepsilon_1, 1)\} \frac{n_1}{n}^{1/\alpha_1} \leq x}{n^{1/\alpha_1}} / \frac{1}{n} \sum_{i=1}^{n} 1\{Y_i = 1\} = \frac{n_1}{n}) dP^Y(j_1, \ldots, j_n)(\#eq:6)$$

Since $\frac{1}{n} \sum_{i=1}^{n} 1\{Y_i = 1\} \underset{\text{n}}{\overset{\text{a.s.}}{\longrightarrow}} b_1$ and by Dominated Convergence Theorem the limit in (#eq:6) equals to

$$\int_0^1 \lim_n P(\frac{max\{f(\varepsilon_1, 1), \ldots, f(b_1 n, 1)\}}{(b_1 n)^{1/\alpha_1}} (b_1)^{1/\alpha_1} \leq x/b_1) dP^{b_1}(u)(\#eq:7)$$

$\frac{max\{f(\varepsilon_1,1),\ldots,f(b_1 n,1)\}}{(b_1 n)^{1/\alpha_1}} \underset{\text{n}}{\overset{\text{w}}{\longrightarrow}} \Phi_{\alpha_1}$

Therefore if $I(\infty)$ is trivial, b_1 is deterministic and (#eq:7) equals to $P(b_1^{1/\alpha_1} \Gamma \leq X)$, with $\Gamma \sim \Phi_{\alpha_1}$ and part a) of Theorem 1 follows.

If b_1 is random using (H2) then part b) of Theorem 1 follows.

$\underline{Remark}$:

It is clear that (H2) may be removed, leading in (b) to an integral with respect to the distribution of b_1 instead of a sum.

$\underline{Example 2}$

Following the ideas of Example 1, consider $\sigma(1), \sigma(2)$ independent such that $P(\sigma(1) = 1) = \delta, P(\sigma(1) = 2)1 - \delta$, $P(\sigma(2) = 1) = \eta, P(\sigma(2) = 2)1 - \eta, 0 < \delta < 1, 0 < \eta < 1$ and $\delta \neq \eta$

Taking $\sigma_1(1), \sigma_1(2), \ldots, \sigma_n(1), \sigma_n(2), \ldots$ a sequence of independent copies of $(\sigma(1), \sigma(2))$ it turns out that if$U$ is a fixed random variable such that $P(U = 1) = p, P(U = 2) = 1 - p, 0 < p < 1$ then if $Y_i = \sigma_i(U)$, we have that $Y_1, \ldots Y_n, \ldots$ fulfills (H1),(H2) with b_1, b_2 random variables such that

$$b_1 = \begin{cases} \delta \text{ if } U = 1 \\ \eta \text{ if } U = 2 \end{cases}$$

and

$$b_2 = \begin{cases} 1 - \delta \text{ if } U = 1 \\ 1 - \eta \text{ if } U = 2 \end{cases}$$

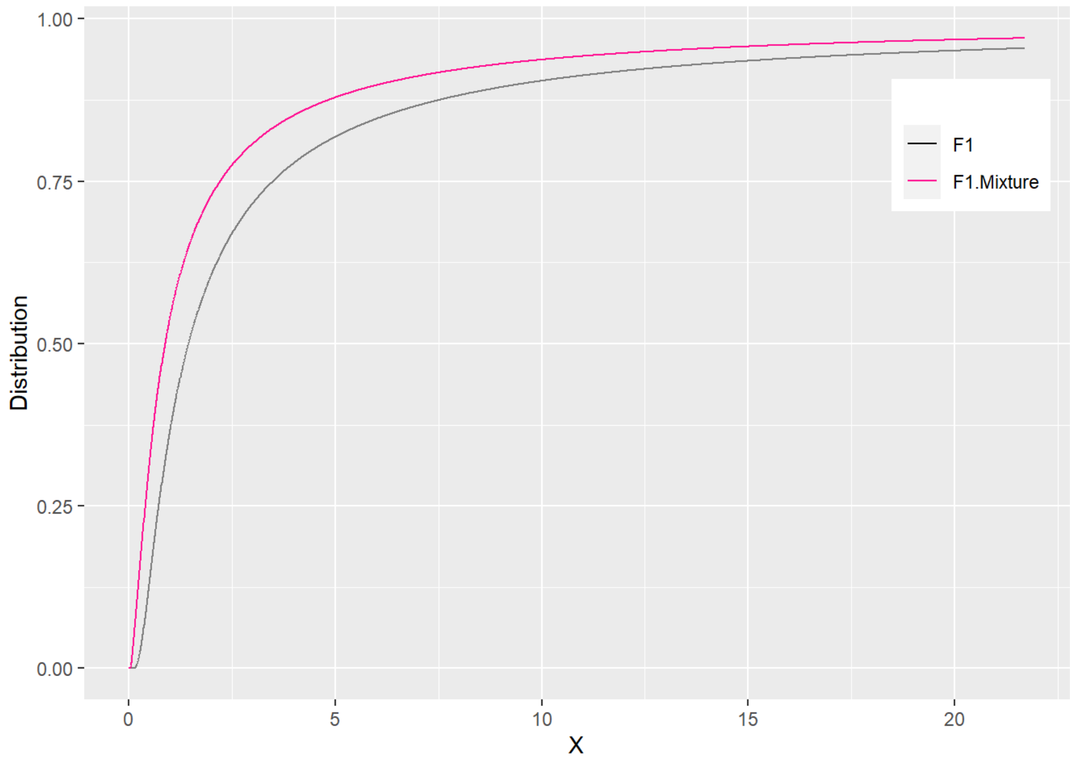
Thus if we assume $0 < \alpha_1 < \alpha_2$ and consider two independent sequences $V_1^{(1)}, \ldots, V_n^{(1)}, \ldots, iid \sim F^{(1)}, V_1^{(2)}, \ldots, V_n^{(2)}, \ldots, iid \sim F^{(2)}$ with $F^{(i)} \in MDA(\Phi_{\alpha_i})$ and we set:

a. If $\sigma_i(U) = 1, X_i = V_i^{(1)}$
b. If $\sigma_i(U) = 2, X_i = V_i^{(2)}$

then, by Corollary 1,

$\frac{max(X_1,\ldots,X_n)}{n^{1/\alpha_1}} \xrightarrow[n]{w} p\Phi_{\alpha_1}\frac{\cdot}{\delta^{1/\alpha_1}} + (1-p)\Phi_{\alpha_1}(\frac{\cdot}{\delta^{1/\alpha_1}})$, a mixture of Frechet distributions.

Figure 1 shows the difference between a Frechet model with $\alpha_1 = 1$ (F1), and a mixture with the same α_1 (F1.Mixture), $p = 0.25$; $\delta = 0.20$; $\eta = 0.80$.



Even for such moderated values, mixtures models is less pesimistic , and for instance considering the value $u = 1.90$ for mixture model the probability of exceeding u is 0.283 and its return time 2.444 (30.9% more frequent). Off course with other settings of parameters differences may be stronger and on the contrary sense (underestimation of simpler models). The aim here is to show than even in not so-complex situations, ignoring the real complexity may lead to significative errors.

$Example3$

We will now simulate data leading to the mixture model of the preceding example and see if even in such a case where differences with a simple Frechet model it is not drammatical, we are able to show which is the most suitable model. If we consider a distribution function of the form:

$$F(x)^{(\alpha)} = 1 - frac1(1+x)^{\alpha}\ \forall x > 0, (\#eq:8)$$

then $F1 - (x)^{(\alpha)} = frac1(1+x)^{\alpha} = \frac{1}{x^{\alpha}}L(x)$, with $L(x) = (\frac{x}{1+x})^{\alpha}$ which is a slowly varying function and therefore a $F^{(\alpha)} \in MDA(\Phi_{\alpha})$.

Its inverse function is

$$F^{(\alpha)-1}(y) = \frac{1}{(1-y)^{1/d}} - 1\ \forall y \in (0,1)(\#eq:9)$$

Consider n large (say $n = 500$) and N very large (say $N = 2000$) and consider $(\sigma(1), \sigma(2))$ a random variable such that $\sigma(1), \sigma(2)$ independent

$P(\sigma(1) = 1) = \delta, P(\sigma(1) = 2) = 1 - \delta\ P(\sigma(2) = 1) = \eta, P(\sigma(2) = 2) = 1 - \eta$, with $0 < \delta < 1, 0 < \eta < 1, \delta \neq \eta$

For instance, we shall take $\delta = 0.20; \eta = 0.80$

Consider $(\sigma_{i,j}(1), \sigma_{i,j}(2))$ (with $1 \leq i \leq N$ and $1 \leq j \leq n$) a matrix of independent copies of $(\sigma(1), \sigma(2))$. As seen in Example 1, if for each i, j we define $Y_{i,j} = \sigma_{i,j}U_j$

Consider $U_1, \ldots, U_N iid$ such that $P(U_j = 1) = p, P(U_j = 2) = 1 - p$ and take $p = 0.25$.

then, for each fixed $i\ Y_{i,1}; \ldots; Y_{i,n}$ fulfills (H1),(H2)

If we consider two independents iid matrix $\varepsilon^{(1)}_{i,j}$ (with $1 \leq i \leq N$ and $1 \leq j \leq n$) and $\varepsilon^{(2)}_{i,j}$ (with $1 \leq i \leq N$ and $1 \leq j \leq n$) such that $\varepsilon^{(1)}_{i,j} \sim F^{(1)}$ (for $\alpha = 1$), $\varepsilon^{(2)}_{i,j} \sim F^{(2)}$ (for $\alpha = 2$)

and we finally set for each $i = 1, \ldots, N$,

$$X_{i,j} = \begin{cases} \varepsilon^{(1)}_{i,j} \text{ if } \sigma_{i,j}(U_i) = 1 \\ \varepsilon^{(2)}_{i,j} \text{ if } \sigma_{i,j}(U_i) = 2 \end{cases}$$

then by Corollary 1, $M_i = \frac{max\{X_{i,1}, \ldots, X_{i,n}\}}{n}$ must be close toe distribution $MF(x) = 0.25\Phi_1\frac{x}{0.2} + 0.75\Phi_1\frac{x}{0.8}$

We will now compute the historam of $M_1, \ldots, M_N$ and see if it fits better to MF or to a classical Φ_1 model.

Conclusions

References

[1]. Aspirot, L., Bertin, K., & Perera, G. (2009). Asymptotic normality of the Nadaraya-Watson estimator for nonstationary functional data and applications to telecommunications. Journal of Nonparametric Statistics, 21(5), 535-551.

[2]. Batt, Ryan D.; Carpenter, Stephen R.; Ives. Anthony R. (2017) Extreme events in lake ecosystem time series. Limnology and Oceanography Letters 2, 63-69. doi: 10.1002/lol2.10037

[3]. Bellanger, L; Perera, G. (2003). Compound Poisson limit theorems for high-level exceedances of some non-stationary processes. Bernoulli Vol 9, No.3, 497-515.

[4]. Bousquet, Nicolas; Bernardara, Pietro (Editors) (2018). Extreme Value Theory with Applications tp Natural Hazards: From Statistical Theory to Industrial Practice, Springer, . ISBN 978-3-030-74941-5

[5]. Embrechts, P.; Kluppelberg, C.; Mikosch, T. (2012).Modelling Extremal Events for Insurance and Finance (9th printing). Springer. ISBN 978-3-642-08242-9

[6]. Fisher, R.A. and Tippett, L.H.C. (1928) Limiting forms of the frequency distributionof the largest or smallest member of a sample. Proc. Cambridge Philos. Soc. 24, 180-190

[7]. Gnedenko, B.V. (1943) Sur la distribution limite du terme d'une serie aleatoire. Ann. Math. 44, 423-453.

[8] Hannesdottir, Asta; Kelly, Mark; Dimitrov, Nkolay (2019). Extreme wind fluctuations:joint statistics, extreme tubulence, and impact on wind turbine loads. Wind Ener. Sci. 4, 325-342.

[9]. Jimenez, E; Cabanas, B. & Lefebvre, G. (Editors) (2015). Environment, Energy and Climate Change I: Environmental Chemistry of Pollutants and Waste. Springer. ISBN 978-3-319-12906-8

[10]. Katz, Richard W.; Brush, Grace S.; Parlange, Grace S.. (2005). Statistics of extremes: modeling ecological disturbances..Ecology, 86(5), pp. 1124-1134.

[11] Knox, J., 1993. Large increase in flood magnitude in response to modest changes in climate, Nature 361, 430-432.

[12]. Leadbetter, M.R. (1983) Extremes and local dependence of stationary sequences. Z. Wahrscheinlichkeitstheorie verw. Gebiete 65, 291-306

[13]. Leadbetter, M.R., Lindgren, G. and Rootzen, H. (1983) Extremes and Related Properties of Random Sequences and Processes. Springer, Berlin

[14]. Leadbetter, M.R. and Rootzen, H. (1988) Extremal theory for stochastic processes. Ann. Probab. 16, 431-478.

[15]. Mostafa, Hany & Roushdi, Mahmoud & Kheireldin, Khaled. (2016). Statistical Analysis of Rainfall Change over the Blue Nile Basin. In 18th International Conference on Environment and Climate Change (ICECC 2016).

[16]. Perera, G. (2002). Irregular sets and central limit theorems. Bernoulli 8, 627-642.

[17]. Reiss, Rolf-Dieter; Thomas, Michael (2007). Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhauser. ISBN 978-3-7643-7399-3

[18]. Rootzen, H. (1978) Extremes of moving averages of stable processes. Ann. Probab. 6, 847-869.

[19]. Rootzen, H. (1986) Extreme value theory for moving average processes. Ann. Probab. 14, 612-652.

[20]. Rootzen, H. (1988) Maxima and exceedances of stationary Markov chains. Adv. Appl. Probab. 20, 371-390.

[21]. Rootzen, H., Leadbetter, M.R. and de Haan, L. (1998) On the distribution of tail array sums for strongly mixing stationary sequences. Ann. Appl. Probab. 8, 868-885.

[22]. Rootzen, H. and Tajvidi, N. (1997) Extreme value statistics and wind storm losses: a case study. Scand. Actuar. J., 70-94.

[23].Willems P. (1998). Hydrological applications of extreme value analysis. In: Hydrology in a changing environment, H. Wheater and C. Kirby (ed.), John Wiley & Sons, Chichester, vol. III, 15-25.